

AD-A176 500

STRUCTURE AND FUNCTIONAL STUDIES OF DEN-2 VIRUS GENOME  
(U) KANSAS UNIV MEDICAL CENTER KANSAS CITY DEPT OF  
BIOCHEMISTRY R K PRADHANABHAN 01 SEP 83  
DAND17-82-C-2051

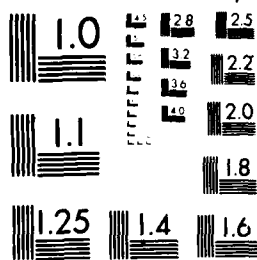
1/1

UNCLASSIFIED

F/G 6/13

NL

END  
DATE  
FILMED  
87  
DTIC



MICROCOPY RESOLUTION TEST CHART  
NATIONAL BUREAU OF STANDARDS-1963-A

AD-A176 500

Report Number AD \_\_\_\_\_

"Structure and Functional Studies of DEN-2 Virus Genome"

Progress Report

by

Radha Krishnan Padmanabhan, Ph.D.

Date: September 1, 1983

DTIC  
ELECTE  
FEB 06 1987  
S D

Supported by

U.S. ARMY MEDICAL RESEARCH AND DEVELOPMENT COMMAND  
Fort Detrick, Frederick, Maryland 21701  
Contract No. DAMD 17-82-C-2051

Prepared by

University of Kansas Medical Center  
Department of Biochemistry  
Kansas City, KS 66103

DTIC DISTRIBUTION STATEMENT

Approved for public release; distribution unlimited.

"The view, opinions and findings contained in this report are those of the author and should not be construed as an official Department of the Army position, policy or decision, unless so designated by other documentation."

FILE COPY

87 2 5 194

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER Two	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) Structure and Functional Studies on Dengue-2 Virus Genome		5. TYPE OF REPORT & PERIOD COVERED Progress Report 1 Sep 82 - 31 Aug 83
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s)  Radha Krishnan Padmanabhan		8. CONTRACT OR GRANT NUMBER(s)  DAMD17-82-C-2051
9. PERFORMING ORGANIZATION NAME AND ADDRESS University of Kansas Medical Center Department of Biochemistry Rainbow Blvd. at 39th St., Kansas City, KS 66103		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS U.S. Army Medical Research Development Command Fort Detrick Frederick, Maryland 21701		12. REPORT DATE 1 Sep 1983
		13. NUMBER OF PAGES
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office) Office of Research Management Walter Reed Army Research Institute Washington, D.C. 20012		15. SECURITY CLASS. (of this report) unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE
16. DISTRIBUTION STATEMENT (of this Report)		
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)		
18. SUPPLEMENTARY NOTES A portion of the work supported by USAMRDC was presented at the annual meeting of the American Society for Biological Chemists, held at San Francisco between June 6 and 9, 1983.		
19. KEY WORDS (Continue on reverse side if necessary and identify by block number) Oligo d(C)-tailed cDNA of DEN-2 RNA annealed to oligo d(G)-tailed pBR322 vector; transformation of <i>E. coli</i> HB101 and characterization of a recombinant DNA from a clone pNK132; cloning of DEN-2 cDNA synthesized from DEN-2 RNA template and random DNA primers using M13mp8 vector/ <i>E. coli</i> JM103 vector/host system; DNA sequence analysis of M13 clones.		
20. ABSTRACT (Continue on reverse side if necessary and identify by block number) 1. We synthesized double-stranded cDNA from DEN-2 RNA template. The heterogeneous mixture of cDNA was tailed with oligo d(C) and cloned at the PstI-cleaved and oligo d(G)-tailed pBR322 DNA. Transformation of <i>E. coli</i> HB101 and screening of colonies gave rise to a clone which contained approximately 400 bp DNA insert. 2. We showed that this 400 bp DNA insert contained DNA sequences complementary to DEN-2 RNA. The evidence came from the experiment in which the nicked 400 bp DNA fragment was shown to serve as a primer for cDNA synthesis by reverse transcriptase.		

on the non-polyadenylated DEN-2 RNA. The size analysis of cDNA synthesized was similar to that obtained when oligo(A)<sub>10</sub> tailed DEN-2 RNA as template and oligo d(T) as primer were used for cDNA synthesis.

3. The 400 bp cDNA insert from DEN-2 RNA was found to contain a site for TaqI restriction enzyme, and we have sequenced the TaqI-B fragment of this insert.
4. We have synthesized cDNA from DEN-2 RNA using random primers, 6 to 8 nucleotides long, which were the end products of pancreatic DNase I digestion of calf thymus DNA. The cDNA synthesized from random primers contained blunt-end at one and Sall site at the other terminus. For cloning the cDNA, M13mp8, cleaved with SmaI and Sall was used as vector and E. coli JM103 as host.
5. We obtained about 500 M13 recombinant plaques, of which only 230 were screened by size analysis of an agarose gel and by Southern hybridization to labeled cDNA probes synthesized DEN-2 RNA. About 60% of the clones were found to contain inserts greater than 300 bases and some up to 1000 bases.
6. DNA sequence analysis of these M13 clones has begun using Sanger's dideoxy chain termination method. We report the DNA sequence data from 8 clones.



Accession For	
NTIS CRA&I	<input checked="" type="checkbox"/>
DTIC TAB	<input type="checkbox"/>
Unannounced	<input type="checkbox"/>
Justification	
By	
Distribution /	
Availability Codes	
Dist	Avail and/or Special
A-1	

## TABLE OF CONTENTS

	<u>Page Number</u>
1. Cover	1
2. Report Documentation Page	2
3. Table of Contents	4
4. Introduction	4
5. Progress Report	4-36
6. Literature Cited	37
7. Personnel Supported During Sept. 1, 1982-August 31, 1983	38

## INTRODUCTION

### PROGRESS REPORT (September 1, 1982-August 31, 1983)

The progress made in the project can be grouped in three parts focussing on the central goal of identifying and characterizing the gene for Dengue virus glycoprotein antigen E.

1. Cloning of complementary DNA (cDNA) copy of Dengue viral RNA into E. coli HB101/pBR322 host/vector system and characterization of a clone that contained a 400 bp insert complementary to Dengue RNA will be described.
2. Construction of a cDNA library using Dengue RNA as template and a random mixture of short oligodeoxy nucleotides (6-8 nucleotides long) as primers for reverse transcriptase will be described. The double-stranded cDNA molecules that were synthesized in this approach were cloned using M13mp8 vector and E. coli JM103 as host. Several positive recombinant bacteriophage plaques were picked based on the insertional inactivation of  $\beta$ -galactosidase marker gene present in the vector. Details regarding further screening of these positive plaques by size analysis of the recombinant DNA and by Southern hybridization will be given.
3. DNA sequence analysis of some of these clones will be described.

1. Cloning of cDNA copy of DEN-2 RNA using E. coli HB101/pBR322 host vector system

- a. Background to Methods

DEN-2 RNA has no polyA tract at its 3' end. In order to synthesize the cDNA of an RNA molecule from its 3' end, one approach involves the use of oligo d(T) as primer for reverse transcriptase. This requires the presence of polyA (or oligoA) at the 3' end of the RNA. I proposed two approaches for the addition of polyA to the 3' end: (1) E. coli polyA polymerase which catalyzes the polymerization of AMP residues (from precursor ATP) to the 3' end of an RNA was used. The length of polyA tail synthesized by E. coli polyA polymerase varies from 300 to several thousand residues, depending on the time of incubation. (2) Alternatively, T4 RNA ligase which can join short oligoA residues to the 3'-OH of DEN-2 RNA was also used. T4 RNA ligase catalyzes the formation of an inter-nucleotide bond between a 5' terminal phosphate and a 3' terminal hydroxyl of oligo- or polyribonucleotides with the accompanying hydrolysis of ATP (England and Uhlenbeck, 1978). RNA ligase was used to add 5'-(<sup>32</sup>P)-labeled oligo(A)<sub>10</sub> to the 3' end of DEN-2 RNA as described in 1982 ANNUAL REPORT.

DEN-2 RNA containing a polyA tract at its 3' end added by RNA ligase was used for cDNA synthesis using oligo d(T) as primer catalyzed by reverse transcriptase. The experimental conditions for cDNA synthesis and its size analysis by alkaline agarose gel (Fig. 5 in 1982 ANNUAL REPORT) were already described. In this report, experimental details of double-stranded cDNA synthesis, insertion into oligo d(G)-tailed pBR322 vector, transformation of E. coli HB101, selection of recombinant clones by sensitivity or resistance to antibiotics, further characterization of recombinant clones by restriction enzyme analysis, hybridization to DEN-2 RNA, large-scale purification of a recombinant DNA plasmid and partial DNA sequence analysis of the 400 bp insert in this plasmid will be described.

## **b. Experimental Details**

The strategy used for cDNA synthesis using DEN-2 RNA template, oligo d(T) primer and reverse transcriptase is outlined in Fig. 1. We also showed that if oligo (A)<sub>10</sub> was not added to the 3' end of DEN-2 RNA in Step 1, then DEN-2 RNA did not serve as template in oligo d(T)-primed-reverse transcriptase reaction for cDNA synthesis. This observation suggested that virion RNA is non-polyadenylated and independently confirmed those made by Blair and Schmaljohn (1979) that virion RNA (from Japanese Encephalitis) did not bind to oligo d(T)-cellulose column chromatography.

Experimental details of Steps 1 and 2 in Fig. 1 are included in 1982 ANNUAL REPORT. The size of cDNA analyzed by alkaline agarose gel electrophoresis (Fig. 5 in 1982 ANNUAL REPORT) is heterogeneous and varied from 400 to >5000 nucleotides long. We believe the reason for this heterogeneity is a high degree of secondary structure present in the viral RNA. Hence, DEN-2 RNA may be a poor template for efficient cDNA synthesis by reverse transcriptase. There is a precedent for this structural feature in viral RNA molecules. Goelet et al. (1982) reported that tobacco mosaic virus RNA proved to be a poor template for the synthesis of cDNA longer than 2000 nucleotides and the efficiency of second-strand synthesis was highly sequence dependent. In order to overcome this problem, the viral RNA has to be denatured prior to cDNA synthesis. Methyl mercuric hydroxide can be used to denature RNA, and excess methyl mercuric ions can be complexed by the addition of  $\beta$ -mercaptoethanol. We are currently testing this method for the synthesis of long cDNA of DEN-2 RNA.

### **Step 3. Oligo d(C)-tailing of cDNA of DEN-2 RNA**

The reaction mixture for the addition of oligo d(C)-tail to the 3' end of cDNA contained, in a total volume of 20  $\mu$ l, 100 mM K-cacodylate, 1 mM CoCl<sub>2</sub> and 0.2 mM dithiothreitol, 1.9 nmoles of dCTP-(<sup>3</sup>H), 2  $\mu$ l of terminal transferase (Lot #730-15 P.L. and 30 units). The reaction was incubated at 37°C, and aliquots at 0' and 15' were withdrawn and analyzed for acid-insoluble radioactivity incorporated into the 3' ends of cDNA.

0 min =	1,036 cpm
15 min =	51,797 cpm

The reaction was stopped by the addition of 2  $\mu$ l of 0.25 M EDTA and extracted with Phenol/CHCl<sub>3</sub>/isoamylalcohol mixture, followed by ether extraction and ethanol precipitation. The ethanol precipitate of cDNA was collected by centrifugation at 12,000 rpm for 1 hr, and the pellet, after drying in the desiccator briefly, was dissolved in 8  $\mu$ l of sterile water.

### **Step 4. Second-strand synthesis of cDNA**

The reaction mixture for second strand synthesis consists, in a final volume of 20  $\mu$ l, HEPES buffer pH 6.9, 10 mM MgCl<sub>2</sub>, 2.5 mM dithiothreitol, 70 mM KCl, 500  $\mu$ M dNTP, 35 pmoles of d(G)<sub>12-18</sub> and 5 units of *E. coli* DNA polymerase I (Klenow's fragment). Incubation was carried out at 15°C for 4 hr. In order to complete the second-strand synthesis of cDNA, reverse transcriptase was also used subsequent to DNA polymerase I reaction. To the above reaction mixture, the following components were added: Tris-HCl (pH 8.3) (100 mM), KCl (0.14 M), MgCl<sub>2</sub> (10 mM), 2-mercaptoethanol (28 mM), dNTP (1 mM each) and reverse transcriptase (20 units). Incubation was

continued at 42° for 1 hr. The reaction was stopped by the addition of 4  $\mu$ l of 0.25 M EDTA.

The completion of second-strand synthesis of cDNA was measured as % of cDNA molecules that became resistant to single-strand specific nuclease S<sub>1</sub>.

S<sub>1</sub> nuclease digestion was carried out using a small aliquot of cDNA in a reaction mixture containing cDNA, NaCl (20 mM), sodium acetate, pH 4.5 (50 mM), ZnSO<sub>4</sub> (1 mM), glycerol (0.5%) and S<sub>1</sub> nuclease (10 units). The incubation was done at 37° for 30 min. The radioactivity rendered acid-soluble was measured by precipitation with trichloroacetic acid. The cDNA molecules were quantitatively resistant to S<sub>1</sub> nuclease, indicating that the second-strand synthesis was complete. The bulk of the reaction mixture after second strand synthesis was extracted with phenol/CHCl<sub>3</sub>/isoamylalcohol mixture, and the cDNA was separated from dNTP by Sephadex G-75 column chromatography. The peaks of radioactivity excluded from the column were pooled and precipitated by the addition of two volumes of ethanol.

#### Step 5. Oligo d(C)-tailing of double-stranded cDNA using terminal transferase

The reaction conditions were same as in Step 3 except that ( $\alpha$ -<sup>32</sup>P)-labeled dCTP (10  $\mu$  Ci) and unlabeled dCTP (1.5 nmoles) were added in a final volume of 58  $\mu$ l. The reaction was initiated by the addition of 30 units (in 2  $\mu$ l) of terminal transferase and by incubation at 37°C for 5 min. The acid-precipitable radioactivity was measured using 1  $\mu$ l aliquots at 0 min (243 cpm) and at the end of 5 min (2954 cpm) of incubation. It was calculated from the radioactivity incorporated to the 3' ends of ds-cDNA that approximately 13 residues of olig d(C) were added to each 3' end. The reaction was stopped and the cDNA was purified as described in Step 4.

#### Step 6. PstI-cleavage of pBR322 DNA

pBR322 DNA vector was prepared on a large scale (1 liter of growth medium) from *E. coli* strain #400, using established procedures (Maniatis *et al.*, 1982). We obtained 900  $\mu$ g of pBR322 DNA from 1 liter culture of *E. coli* cells.

pBR322 DNA (10  $\mu$ g) was digested with PstI in an incubation mixture containing 20 mM Tris-HCl, pH 7.5, 10 mM MgCl<sub>2</sub>, 1 mM dithiothreitol and 10 units of enzyme. The incubation was carried out at 37° for 2 hr, and the completion of digestion was checked by agarose gel electrophoresis. The reaction was stopped by the addition of EDTA (15 mM), and the DNA was extracted with phenol and precipitated with ethanol.

#### Step 7. Oligo d(G)-tailing of PstI-linear of pBR322 DNA using terminal transferase

The conditions for tailing reaction were the same as described for cDNA except that dGTP-(<sup>3</sup>H) (specific activity 11.2 Ci/mmol; 180  $\mu$ M) was used in 50  $\mu$ l. The incubation was carried out at 37° for 7 min. From the amount of incorporation (0 min = 1651 cpm; 7 min = 22,320 cpm), it was calculated that about 12 dG residues were incorporated per 3' end. The reaction was stopped and the DNA was purified by Sephadex G-75 column chromatography as described above. The peak fractions of DNA were pooled and precipitated with ethanol. The pellet collected after centrifugation and drying briefly in the desiccator was dissolved in 100  $\mu$ l of annealing buffer (10 mM Tris-HCl, pH 7.5, 1 mM EDTA and 0.1 M NaCl).

**Step 8. Transformation of *E. coli* HB101 with annealed mixture of oligo d(G)-tailed pBR322 and oligo d(C)-tailed cDNA**

*E. coli* HB101 cells were made competent for transformation as follows. An overnight culture of *E. coli* HB101 in 5 ml of Luria-Bertani (LB) broth was used to inoculate 25 ml of LB broth (1 to 100) in a 250 ml Erlenmeyer flask which was vigorously shaken at 37°C. When the O.D.<sub>550 nm</sub> reached about 0.5, the flask was chilled in ice and the cells were pelleted by centrifugation at 8000 rpm in a Sorvall at 4°C for 5 min. The supernatant was discarded, and the cells were resuspended in 10 ml of transformation buffer (50 mM CaCl<sub>2</sub>, 250 mM KCl, 5 mM Tris-HCl, pH 7.5, and 5 mM MgCl<sub>2</sub>). The cells were resuspended gently by pipeting the solution a few times, and the cells were incubated in ice for 20 min. The cells were pelleted once again by centrifugation and resuspended in 2 ml of transformation buffer. The cells were kept at 4°C overnight.

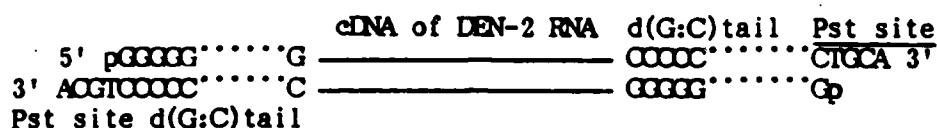
Oligo d(G)-tailed pBR322 (100 ng) was mixed with DEN-2 cDNA (✓ 10 ng) in a total volume of 100 µl of annealing buffer. The mixture was heated at 65°C for 5 min and incubated at 57°C for 2 hr. Competent cells (0.2 ml) of *E. coli* were added to the annealed mixture as well as to appropriate control DNAs (such as uncleaved pBR322 DNA, oligo d(G)-tailed DNA alone). The mixture was incubated in ice for 25 min and then at 37°C for 5 min. Then, LB broth (1 ml) was added to each tube and the tubes were incubated at 37°C for 1 hr. Aliquots (0.3 ml) were plated on LB agar plates containing tetracyclin (20 µg/ml). The agar plates were incubated at 37°C overnight for the transformed colonies to grow. A total of 226 tetracyclin-resistant colonies were obtained. These colonies were then tested for ampicillin (Ap<sup>R</sup>) sensitivity by plating them on agar plates containing ampicillin (50 µg/ml). A total of 50 colonies was found to be Ap sensitive. Twenty Ap sensitive and Tc<sup>R</sup> colonies were screened first for the size of the DEN-2 cDNA insert. The plasmid DNA was isolated from each clone by a miniplasmid isolation method. The clones were grown overnight in 2 ml of LB broth containing 20 µg/ml tetracyclin. The cells (1.5 ml) were centrifuged in an Eppendorf tube for 1 min at 4°C. The supernatant was discarded, and the pellet was resuspended in a mixture containing 50 mM glucose, 10 mM EDTA, 25 mM Tris-HCl (pH 8.0) and 4 mg/ml of lysozyme and incubated for 5 min at room temperature. To this, a 200 µl aliquot of 0.2 N NaOH containing 1% SDS was added, and the contents were mixed gently 3 times. After incubation of this mixture in ice for 5 min, 150 µl of ice-cold solution of K-acetate (pH 5.0) (3 M in K<sup>+</sup> and 5 M in Ac<sup>-</sup>) was added. After mixing, the tube was incubated in ice for 5 min and then centrifuged in an Eppendorf Microfuge for 5 min at 4°C. The supernatant was transferred to another Eppendorf tube. The mixture was extracted with phenol/CHCl<sub>3</sub>/isoamylalcohol mixture and precipitated with ethanol. The pellet of DNA was washed with 70% ethanol and was dried. The dry pellet was dissolved in 50 µl of TE (10 mM Tris-1 mM EDTA) buffer.

The plasmid DNA isolated from 1.5 ml culture of recombinant clone was analyzed by PstI cleavage and agarose gel electrophoresis. The results are shown in Fig. 2. Six clones, I-32, I-44, II-14, II-19, II-43 and III-3, were found to contain the largest insert, which was only 500 bp long. Similar analyses of 20 more clones showed that all clones had inserts of similar size (Fig. 3). Nine positive clones based on the size of the insert released subsequent to PstI cleavage were further analyzed by a double restriction enzyme digestion using BamHI and PvuII. pBR322 vector alone gives rise to two DNA fragments, 2670 bp and 1690 bp long. Any cDNA insert cloned at PstI site will increase the size of the 2670 bp fragment and hence decrease its electrophoretic mobility on an agarose gel, depending on the size of the insert. Fig. 4 shows the results of this analysis, which indicated that all 9 clones had cDNA inserts but the size of the

inserts varied from 200-400 bp. Only two clones, I-32 and II-19, had longer inserts and hence were chosen for further study.

Plasmid DNA from clones I-32 and II-19 were isolated by growing a 50 ml culture. PstI digest and analysis on an agarose gel electrophoresis is shown in Fig. 5. The 400 bp insert from I-32 clone was isolated by a preparative gel electrophoresis. The DNA fragment was cut out and purified from the gel using the glass powder method (Pettersson, 1979).

In order to investigate whether this fragment contained the complementary sequences to DEN-2 RNA, the 400 bp fragment was nicked using pancreatic DNase I to generate oligodeoxynucleotides which were then tested in a reverse transcriptase-catalyzed cDNA synthesis on a DEN-2 RNA template. The rationale for nicking the 400 bp fragment is that the fragment had d(G:C) tails on both ends which are not complementary to the viral RNA (see below)



The nick translation of the 400 bp fragment was carried out in a reaction mixture (50  $\mu$  l) containing 0.1  $\mu$  g of DNA, 1 mM each of dNTP, 50 mM Tris-HCl, pH 7.5, 5 mM MgCl<sub>2</sub>, 50  $\mu$  g/ml bovine serum albumin, 1 mM dithiothreitol and diluted DNase I (1  $\mu$  l of 333  $\mu$  g/ml solution diluted to 1 ml of 10 mM Tris-HCl, pH 7.5). The reaction mixture was heated in a boiling water bath for 5 min and quick cooled in ice. This mixture was used as such for a reverse transcriptase-catalyzed cDNA synthesis on a nonpolyadenylated DEN-2 RNA. The reverse transcriptase reaction contained 0.6  $\mu$  g of DEN-2 RNA (0.14 pmole), 0.17 pmole of nicked 400 bp fragment, 25  $\mu$  Ci of dCT<sup>32</sup>P, 2.5 nmole of unlabeled dCTP and 13.3 nmole of each of the other dNTP. Other conditions of the reaction were similar to those mentioned earlier. Under these conditions, per mole of DEN-2 RNA, 900 moles of dCMP residues were incorporated into cDNA. These results indicated that the nicked 400 bp DNA insert contained sequences complementary to DEN-2 RNA because it could serve as primer for cDNA synthesis by reverse transcriptase.

Analysis of the size of the cDNA was carried out by alkaline agarose gel. The autoradiography of the gel is shown in Fig. 5. The size of cDNA was heterogeneous from 400 to about 5000 nucleotides long, although predominant species were 1000 nucleotides long. This heterogeneity in the cDNA molecules might be due to the interference from the secondary structure of the viral RNA as discussed earlier. Denaturation of the viral RNA by methyl mercuric hydroxide followed by cDNA synthesis might give rise to long cDNA molecules and this remains to be tested.

We have recently used the oligonucleotides generated from the nicked 400 bp DEN-cDNA as primers for reverse transcriptase on the nonpolyadenylated (native) DEN-2 RNA as template. The heterogeneous cDNA that was synthesized was tailed with oligo-d(C) (14 residues were added per 3' end). Second-strand synthesis was primed by oligo-d(G) (step 4 in Fig. 1). Sall-linker was ligated to the end that was primed by oligomers from the 400 bp fragment. The cDNA were digested with Sall and fractionated according to size by Sepharose-4B column chromatography.

The fractions containing cDNA larger than 2 kb were pooled and precipitated with ethanol. This fraction of cDNA remains to be cloned into pUC8 and pUC9/*E. coli* JM83 vector/host system (see the Renewal Application).

2. Shot-gun cloning of cDNA synthesized on DEN-2 RNA template using random primers into M13mp8

a. Background to Methods

One approach to analyze the organization of a viral genome and localization of the genes coding for viral-specific structural and non-structural proteins is to determine the complete nucleotide sequence from which the amino acid sequence of the protein can be deduced. These data can be compared with those obtained from microsequence analysis of specific antigens coded by the virus in order to unambiguously locate the respective genes on the viral genome. It has become possible to determine the complete nucleotide sequence of large viral genomes (even as large as 35 kb long adenovirus type 2 DNA and 50 kb long  $\lambda$  DNA) due to the development of rapid methodology for sequencing DNA. Complete sequence analysis of poliovirus RNA (7433 nucleotides long) and tobacco mosaic virus RNA (6000 nucleotides long) have been determined, and their genes encoding the structural and non-structural proteins have been elucidated. The strategy generally used for such large sequencing projects is a random approach. The most rapid random approach is that of Sanger and his colleagues (Sanger et al., 1977, 1980), who sequence target DNA cloned into a M13 vector using the dideoxy chain termination method. In this approach, randomly cleaved target DNA is molecularly cloned adjacent to a site where an universal primer hybridizes and which can be elongated to different extents by DNA polymerase I (Klenow's fragment) in the presence of normal dNTP as well as one of the four specific chain terminator dideoxy dNTP (ddNTP). By using four different ddNTP in four separate reactions as chain terminators, labeled (-) strand of M13 is synthesized from the 3'-OH of the universal primer and terminated at each complementary nucleotide. By fractionating these fragments generated in four separate dideoxy chain terminator reactions on a sequencing gel and autoradiography, one can conveniently acquire about 250 nucleotides of sequence data per sequencing experiment (Fig. 7).

The filamentous single-stranded DNA phages of male specific *E. coli* offer numerous advantages as cloning vehicles. Both large amounts of replicative form I DNA (100-200 copies per cell) and single-stranded DNA can be isolated. Due to asymmetric DNA synthesis only one strand of the cloned DNA is synthesized in large amounts and packaged, depending on the orientation of insertion, into replicative form of the phage vector. This results in the purification of that strand. The high titer of the bacteriophage permits the isolation of large amounts of any cloned DNA fragment. We used M13mp8 vector (Fig. 8) for cloning the cDNA of DEN-2 RNA. The vector has a 789 bp insertion of *E. coli* DNA that codes for the first 145 amino acids of the  $\beta$ -galactosidase gene. When M13mp8 infects an *E. coli* (e.g., JM103 strain) that contains an appropriate partial deletion of  $\beta$ -galactosidase, the phage-encoded peptide fragment can complement the host fragment ( $\alpha$ -complementation), resulting in blue plaques on indicator medium in the presence of the inducer isopropylthiogalactoside (IPTG) and chromogenic substrate (5-bromo-4-chloro-3-indolyl- $\beta$ -galactopyranoside or X-gal). The vector has a multiple cloning site (MCS) containing several restriction endonuclease cleavage sites in close proximity. When a foreign DNA fragment is cloned into the RF, the resultant phage makes a colorless plaque after transfection due to insertional inactivation of  $\beta$ -galactosidase gene on the phage vector. This feature allows easy screening of recombinant phages.

The random approach and M13 cloning have been used for sequence analysis of cDNA copies of large RNA genomes such as TMV RNA (Goelet *et al.*, 1982). Cloned poliovirus type 3 cDNA was subcloned into the mp7, 8 and 9 derivatives of bacteriophage M13 and sequenced by the dideoxy method (Cann *et al.*, 1983). An optimal library for DNA sequencing should contain clones of fragments having overlapping sequences and clones having equal representation from every part of the DNA molecule. In the case of TMV RNA, the overlapping cDNA molecules representing the entire genome were synthesized using a primer "cocktail" consisting of a mixture of 4 to 7 residue long oligodeoxynucleotides synthesized chemically (Goelet *et al.*, 1982). Another source of primer cocktails for cDNA synthesis of long RNA is the end product of DNase I digestion of calf thymus DNA. This mixture of short oligodeoxynucleotides, 6-8 residues long, has been shown to function as primers for the efficient copying into DNA of many naturally occurring RNA's by purified reverse transcriptase (Taylor *et al.*, 1976).

The DNA sequence data that are gathered by this approach are fed into a computer in order to identify and order each block of sequence data (Staden, 1980). Perhaps 80% of a large sequencing project may be determined fairly rapidly with this approach before increasingly repetitive sequence data are collected. The remaining portion of the genome can be sequenced using specific DNA primers for cDNA synthesis from regions that have already been sequenced.

#### b. Experimental Details

##### i. Preparation of random primer cocktail from calf thymus DNA

Calf thymus DNA (250 mg) was dissolved in 7 ml of 20 mM Tris-HCl, pH 7.4, and 10 mM MgCl<sub>2</sub>. The DNA was incubated with pancreatic DNase I (1 mg) at 37°C for 1 hr. Then, 3 ml of 10% SDS and 1.5 ml of pronase (20 mg/ml) were added, and the incubation was continued at 37°C for 45 min. The solution was extracted with phenol/CHCl<sub>3</sub> mixture, and the aqueous phase was transferred to a siliconized glass tube. The oligodeoxynucleotide mixture was denatured by boiling for 15 min and quick cooled. NaCl was added to a final concentration of 0.1 M. It was then loaded onto a DEAE-cellulose column equilibrated with a solution of 5 mM Tris-HCl, pH 7.4, 0.1 M NaCl and 1 mM EDTA. The column was washed with the same buffer and eluted with 0.3 M NaCl containing 5 mM Tris-HCl, pH 7.4, and 1 mM EDTA. Fractions of 0.5 ml were collected, and fractions 39-50 were pooled and precipitated with ethanol. The ethanol precipitate was centrifuged at 8000 rpm in Sorvall (SS-34 rotor) for 30 min and the pellet was dried in the desiccator. It was then dissolved in 2 ml of TE buffer. The concentration of oligodeoxynucleotides in this solution was 25.35 mg/ml (or 50.7 mg total), which represented 20% of the input DNA.

In order to estimate the size distribution of the DNA primer cocktail, a small aliquot (253  $\mu$ g) was dephosphorylated in a reaction mixture containing 50 mM Tris-HCl, pH 9.0, 1 mM MgCl<sub>2</sub>, 0.1 mM ZnCl<sub>2</sub>, 1 mM spermidine and 10  $\mu$ g of the enzyme. The incubation was carried out at 56°C for 1 hr. The DNA was extracted with phenol/CHCl<sub>3</sub> and precipitated with ethanol. The pellet was dissolved in 25  $\mu$ l of TE buffer. An aliquot (50  $\mu$ g) was labeled with <sup>32</sup>P at the 5' end using polynucleotide kinase and  $\gamma$ -AT<sup>32</sup>P. Synthetic oligodeoxynucleotides 8 and 14 nucleotides long, chemically synthesized using phosphodiester methods of Khorana and coworkers (Padmanabhan *et al.*, 1974; Padmanabhan, 1977), were also labeled at their 5' termini for their use as DNA markers. These labeled oligomers were purified by Sephadex G-75 column chromatography. The size distribution of oligomers in the calf thymus DNA primer cocktail was determined by a 20% polyacrylamide gel electrophoresis using the synthetic DNA

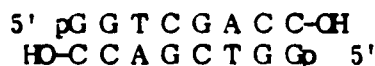
markers of known length. The autoradiography of the gel as shown in Fig. 9 indicates that the primer cocktail contained oligomers 5-6 nucleotides in length.

## ii. cDNA synthesis

This random primer cocktail was used as primers for cDNA synthesis on DEN-2 RNA template. DEN-2 RNA (0.35 pmole) and primer cocktail (250  $\mu$ g) were used for reverse transcriptase-catalyzed cDNA synthesis under the same conditions described earlier, except that actinomycin D (100  $\mu$ g/ml) was included in one half of the reaction mixture. Actinomycin D inhibits the second strand cDNA synthesis by reverse transcriptase, presumably by inhibiting the RNase H activity inherent with the polymerase activity. We found that the size of the cDNA synthesized using random primers was in the range of 100-1400 nucleotides (predominant species being 600 nucleotides) in the absence of actinomycin D, whereas the cDNA synthesized in its presence were found to be slightly longer (up to 2 kb) see Fig. 10). The size of the cDNA molecules synthesized was heterogeneous as expected due to random priming on DEN-2 RNA. The cDNA was purified by Sephadex G-75 column chromatography.

The double-strand cDNA was synthesized by "flip-back" mechanism using *E. coli* DNA polymerase I (see Fig. 11). The reaction mixture for second-strand synthesis contained in 50  $\mu$ l, the cDNA 100 mM HEPES buffer, 10 mM  $MgCl_2$ , 2 mM dithiothreitol, 60 mM KCl, 0.5 mM each of dNTP and 30 units of DNA polymerase I (Klenow's fragment). The incubation was carried out at 15 $^{\circ}$  overnight. The reaction was monitored by sensitivity to  $S_1$  nuclease. An aliquot (1  $\mu$ l) was incubated with  $S_1$  nuclease as described earlier and was found to contain about 80% of  $S_1$  nuclease-resistant radioactivity. Every cDNA molecule will lose some sequences at the hairpin region (Fig. 11). However, since the cDNA synthesized from random primers do overlap, the sequences lost in one cDNA molecule might be present in another cDNA molecule. The second strand synthesis was completed by the addition of reverse transcriptase to the reaction mixture, which was then extracted with phenol and precipitated with ethanol.

The ethanol precipitated DNA was dissolved in 11  $\mu$ l of TE and treated with Klenow's DNA polymerase I to ensure blunt ends at the 3' ends of hair-pin structures (see Fig. 11). The conditions were the same as described above for second-strand synthesis, except that the incubation was at room temperature for 30 min. To this reaction mixture, 5'-phosphorylated Sall linker (1  $\mu$ g, 100 pmole) with the sequence



and, in addition, 25 mM Tris-HCl, pH 7.4, 5 mM  $MgCl_2$ , 5 mM dithiothreitol, 0.25 mM spermidine, 1 mM ATP, 10  $\mu$ g/ml bovine serum albumin, 400 units of T4 DNA ligase and 30 units of T4 RNA ligase. The incubation was carried out at 4 $^{\circ}$ C for 16 hr.

After ligation of Sall linkers to the cDNA, the mixture was treated with  $S_1$  nuclease by adding 8  $\mu$ l of 10X buffer and 10 units of  $S_1$  nuclease. Incubation was carried out at 37 $^{\circ}$  for 30'. The reaction mixture was then extracted with phenol/ $CHCl_3$  and the cDNA was precipitated with ethanol. The cDNA was then digested with Sall to produce 5' protruding Sall sites at one end of all cDNA molecules. The other end was blunt-ended by  $S_1$  nuclease digestion of hair-pin structures.

iii. Size-fractionation of cDNA molecules by Sepharose 4B column chromatography

The cDNA synthesized using random primers on DEN-2 RNA was heterogeneous in size, ranging from 100-1400 base pairs in length. This mixture of cDNA molecules had to be fractionated according to their sizes. We used Sepharose 4B column chromatography to achieve this goal. A column of 1.5 ml bed volume was packed in a pasteur pipette and equilibrated with 10 mM Tris-HCl, pH 7.5, 1 mM EDTA and 0.1 M NaCl. Fractions of 50  $\mu$ l were collected and counted for radioactivity. The sizes of the cDNA in these fractions were determined by agarose gel electrophoresis and autoradiography. Fraction numbers 10, 12, 14, 16, 19, 22 and 24 were used for gel electrophoresis, and the results of autoradiography are shown in Fig. 12. Three pools (Pool I-fractions 7 to 12, Pool II-fractions 13 to 18 and Pool III-fractions 19 to 23) were precipitated with ethanol, and the cDNA was used for transformation of *E. coli* JM103.

iv. Use of *E. coli* JM103/M13mp8 host-vector system for cloning the cDNA of DEN-2 RNA

The multiple cloning sites (MCS) of the vector M13mp8 are shown in Fig. 8. The cDNA molecules of different size classes having asymmetric termini (a blunt and a Sall end) was cloned between SmaI site (blunt) and Sall site in the MCS region of M13mp8. This vector was prepared in large amounts from 500 ml culture of *E. coli* JM103 transfected with M13mp8 vector. The cells were lysed by NaOH-SDS method of Birnboim and Doty (1979), and the replicative form (RFI) of M13mp8 was purified by equilibrium CsCl density gradient centrifugation as described by Maniatis et al. (1982).

M13mp8 DNA (440 ng) was digested with 18 units of SmaI, followed by digestion with Sall (10 units). The mixture was then treated with calf intestinal alkaline phosphatase to dephosphorylate the 5' termini of the vector. The DNA was extracted with phenol/CHCl<sub>3</sub> and precipitated with ethanol. The cDNA from Pool I, II and III (10-20 ng) was ligated to M13mp8 vector (the ratio of cDNA/vector was in the range of 0.25-0.5) in reaction mixture (18  $\mu$ l) containing T4 DNA ligase and T4 RNA ligase as described earlier. The incubation was carried out at 4°C overnight.

Competent cells of *E. coli* JM103 were prepared and were transformed with the ligated mixture. After overnight incubation at 37°C, the plaques were counted. Any blue plaques appearing in the agar plate containing IPTG and the chromogenic substrate, X-gal is due to the background (JM103 transformed by the vector with no insert). Any colorless plaque appearing is considered as positive and picked for isolation of M13 DNA and further characterization. Since the dephosphorylated vector was used for ligation, the background was very low (about 1%). cDNA from Pool I gave rise to about 200 colorless plaques and cDNA (Pool II) gave rise to about 300 plaques. These plaques were picked using sterile toothpicks and grown in 1 ml of 2XYT (yeast extract + tryptone) medium. Freshly grown JM103 cells were also added to the growth medium to increase the titer of M13 phage. After 6-8 hours, the cells were transferred to an Eppendorf centrifuge tube and pelleted. The pellet was resuspended in a minimal medium containing 15% glycerol and stored at -70°C as stocks. The supernatant containing the M13 phage (10  $\mu$ l was saved as stocks) was treated with 200  $\mu$ l of 20% polyethylene glycol (Carbowax 6000) containing 2.5 M NaCl. The solution was kept at room temperature for 30 min. The tubes were centrifuged and the pellet was resuspended in 100  $\mu$ l of TE buffer containing 0.1 M NaCl and extracted with phenol/CHCl<sub>3</sub>, ether and then precipitated with ethanol.

M13 DNA was isolated from a total of 230 colorless plaques for further screening. First, recombinant M13 DNA's (from 140 plaques) were electrophoresed on an agarose gel (1%), and their mobilities were compared with the control M13 DNA with no insert. The mobilities of some of the clones are shown in Fig. 13 (a & b).

These analyses indicated that about 50-60% of the clones migrated slower than the control DNA. Some of these clones had considerably lower mobilities, indicating that they probably contained longer inserts.

The second screening method we used is based on the ability of these clones to hybridize to the ( $^{32}\text{P}$ )-labeled cDNA probes synthesized from DEN-2 RNA using random primers. Since the double-stranded cDNA that was synthesized using random approach and asymmetric cloning sites (SmaI and Sall), the polarity of M13 DNA inserted into M13mp8 vector will be the same as that of RNA. Hence, labeled single-stranded cDNAs synthesized from DEN-2 RNA are expected to hybridize with M13 recombinant clones. The M13 DNA samples were denatured by heating in a 100° water bath for 10 min, followed by chilling in ice. Each sample was treated with 4  $\mu$ l of 1 M NaOH and incubated at room temperature for 20 min. The solution was neutralized by the addition of an equal volume of 1 M NaCl/0.3 M sodium citrate/0.5 M Tris-HCl, pH 8.0, and 4  $\mu$ l of 1 M HCl. The solution was chilled in ice. The samples were spotted on nitrocellulose paper, and the filter was soaked in 6XSSC solution (20XSSC = 175.3 g of NaCl, 88.2 g of sodium citrate in 800 ml H<sub>2</sub>O; the pH was adjusted to 7.0, and the volume was made up to 1 liter) and dried overnight in a 65°C oven. The filters were soaked again in 6XSSC and incubated in a heat-sealable bag with prehybridization solution at 68° in a water bath for 4 hr. The filters were hybridized with ( $^{32}\text{P}$ ) labeled probe in hybridization solution by incubating overnight. The filters were then washed successively with 2XSSC/0.5% SDS, 2XSSC/0.1% SDS and two washings with 0.2XSSC/0.5% SDS to remove the free labeled probe. The filters were then dried at room temperature and autoradiographed at -70°C using Kodak SB-5 X-ray film. Using this technique, only 20 plaques gave positive signals. The reasons for this discrepancy in the number of clones that seemed to have inserts as seen from their mobilities in agarose gel electrophoresis and the number of clones that gave positive signals in hybridization may be that 1) the cDNA molecules synthesized from random primers used as hybridization probes do not represent equally all regions of DEN-2 RNA which served as template; some regions are more efficiently copied than others due to secondary structures in DEN-2 RNA; these underrepresented cDNA molecules used as probes would hybridize to M13 cDNAs to give rise to weak signals due probably to limiting amounts of probe; or that 2) in some cDNA molecules and vector DNA molecules, Sall site might have been blunt-ended due to traces of single-strand specific exonuclease present in the commercial restriction endonucleases (e.g., Sall in this case), blunt-end DNA ligation would have given rise to two possible orientations. Only 50% of these clones would hybridize with the probe. We have evidence for this possibility from the DNA sequence data (see the following section). The results of the dot hybridization experiments are shown in Fig. 14. A total of 32 clones gave positive signals in this screening procedure.

Before these clones were used for DNA sequence analysis, they had to be plaque purified. The colorless plaques that were picked from the original transfection agar plate were grown in JM103 cells, and the phage was isolated by PEG precipitation as described above. These phage stocks were used for plaque purification. First, the stocks were diluted in 10 mM Tris-HCl, pH 7.5, 10<sup>6</sup>-fold. Aliquots of 1  $\mu$ l and 10  $\mu$ l of this diluted phage were mixed with freshly grown JM103 cells (200  $\mu$ l), 30  $\mu$ l of X-gal (2%), 20  $\mu$ l of IPTG (100 mM) and 3 ml of top agar. This mixture was plated on minimal agar plate containing glucose. The plates were incubated at 37°C overnight. Colorless

plaques appeared on a lawn of JM103 cells. A single plaque from this plate was used for isolation of M13 DNA. The M13 DNA purified from a single plaque was treated with 0.2 M NaOH (final concentration) for 2 hr at 37°C to eliminate the background during DNA sequence analysis by Sanger's method. The solution was neutralized and precipitated with ethanol. Gel electrophoresis of plaque purified M13 DNA is shown in Fig. 15.

We were interested to investigate whether the region of DEN-2 RNA contained in the 400 bp DEN-cDNA clone isolated by the strategy illustrated in Fig. 1 is also represented in the M13 cDNA library. The 400 bp DEN-cDNA fragment was nick-translated as described earlier and used as a hybridization probe for screening the M13 clones. A total of 79 M13 clones was applied to two nitrocellulose filters and hybridized to the probe. The results are shown in Fig. 16 a & b. There were six positive clones in our M13 library which contain sequences complementary to the 400 bp clone obtained by a different cloning strategy. These clones most likely represent DEN-2 sequences near the 3' end. Whether the entire 3'-terminal sequences are represented in these clones or not remains to be elucidated. DNA sequence analyses of the 400 bp fragment as well as the corresponding M13 clones are discussed in the next section.

### 3. DNA sequence analysis of DEN-2 cDNA clones

#### a. Background to Methods

There are two rapid methods available for sequencing DNA. 1) Chemical method of Maxam-Gilbert (1977) is based on a partial base-specific chemical cleavage of a single end-labeled DNA fragment, followed by separation of these nested sets of labeled fragments on a denaturing polyacrylamide gel and autoradiography. Although this method is rapid, there is a considerable amount of labor involved in the purification of the cloned DNA fragment (from a plasmid), labeling and restriction endonuclease cleavage followed by fractionation of the single-end labeled fragments. 2) The enzymatic method of Sanger and coworkers (1977, 1980) is based on the use of DNA polymerase I-catalyzed DNA synthesis on a single-stranded DNA template to be sequenced from a primer homology. For single-stranded DNA template, the M13 system is used for cloning the DNA fragment to be sequenced (Sanger *et al.*, 1980). The details of this method are discussed in the Renewal Application.

#### b. Experimental Details

##### i. DNA sequence analysis of the 400 bp DEN-2 cDNA fragment

The plasmid (50 µg) carrying the 400 bp insert of DEN-2 cDNA (pNK132) was digested with PstI (100 units) at 37°C for 60 min. The reaction mixture was heated at 65°C for 10 min and then treated with calf intestinal alkaline phosphatase at 55°C for 60 min to dephosphorylate the 5' ends. The DNA was extracted with phenol/CHCl<sub>3</sub> and ether and precipitated with ethanol. The DNA was dissolved in 80 µl of TE and fractionated by a preparative agarose gel (1%) electrophoresis. The 400 bp fragment was cut out and eluted from the gel using the glass powder method. The DNA, after ethanol precipitation, was dissolved in 8 µl of 0.1 M Tris-HCl, pH 9.5, 5 mM spermidine and 0.5 mM EDTA, heated at 70°C for 5 min and chilled in ice. The 5' ends were labeled with (<sup>32</sup>P) using polynucleotide kinase (20 units) and γ-AT<sup>32</sup>P (sp. activity 3000 Ci/mmol; 60 pmoles) in a total volume of 60 µl. The labeled DNA was purified by Sephadex G-50 column chromatography and precipitated with ethanol. A small aliquot of labeled DNA was digested with AluI, HhaI, HpaII and TaqI enzymes. The digests were electrophoresed on a PAGE (20%) and autoradiographed. The results are shown in

Fig. 17. Only *TaqI* had a site in this 400 bp fragment (the reason for the abnormal mobility of *HhaI* is unknown). Hence, *TaqI* was used for cleaving the bulk of the 5' labeled DNA. *TaqI* digest was fractionated on a PAGE (8%). The two *TaqI* fragments were located on the gel by autoradiography, (see Fig. 18) cut out and eluted from the gel for sequence analysis. Although the two *TaqI* fragments were well separated from each other (A and B in Fig. 18), the large fragment A migrated very close to a contaminant (A") which might have arisen from *E. coli* chromosomal DNA. We have not completed the sequence analysis of A fragment for this reason. However, B fragment could be eluted from the gel as a pure fragment and used for DNA sequence analysis. Autoradiography of Maxam-Gilbert DNA sequencing gel is shown in Fig. 19. The sequence that was read off from this gel and a duplicate experiment are shown in Fig. 20.

ii. DNA sequence analysis of M13 clones of DEN-2 cDNA

Any large sequencing project, as discussed earlier, has been carried out using M13 cloning and Sanger's dideoxy method. We are currently sequencing the M13 clones of DEN-2 cDNA, including those that are complementary to pNK132 plasmid clone. DNA sequence of the *TaqI*-B fragment from the 400 bp insert determined by the Maxam-Gilbert method (Fig. 20) was identical to the corresponding region of an M13 clone (I-102) that hybridized with the 400 bp fragment (Fig. 16 a & b), as shown in Fig. 21. The region underlined is identical to the sequence of *TaqI*-B fragment and, in addition, the dideoxy sequencing of the M13 clone I-102 also gave a portion of the sequence of *TaqI*-A fragment which we have not yet determined by Maxam-Gilbert method. The perfect agreement in DNA sequence analyses of a portion of DEN-2 cDNA by two different methods gives us confidence that both systems are working well in our hands. Partial DNA sequences of other M13 clones are also given in Figs. 22 and 23. As these sequence data accumulate, we are planning to use a computer to compile these data, identify overlapping regions of various clones and arrange them into a long linear sequence of DEN-cDNA.

### Figure Legends

**Figure 1.** Strategy for cloning cDNA from nonpolyadenylated DEN-2 RNA.

**Figure 2.** Analysis of cDNA clones by cleavage with PstI.

a. Ampicillin sensitive, tetracyclin resistant colonies were screened for cDNA inserts by PstI cleavage. Lanes 1-20: plasmid DNA's from clones; group I series - 2, 5, 10, 13, 32, 44; group II series - 4, 5, 7, 14, 19, 20, 36, 43, 44; group III series - 1, 2, 3, 4, 14; Lane 21:  $\lambda$  DNA cleaved with EcoRI and HindIII as marker DNA; Lane 22: uncut pBR322 DNA as control; Lane 23: linear pBR322 DNA as control. Electrophoresis was carried out on a 1% agarose gel at 4 V/cm overnight. The gel was stained with ethidium bromide and the fluorescence was photographed under UV-light.

**Figure 3.** Analysis of cDNA clones by cleavage with PstI.

Screening of more clones which are ampicillin sensitive and tetracyclin resistant was carried out. Lanes 1-20: plasmid DNA's from clones; group III series - 15, 16, 43, 44; group IV series - 1, 24, 25, 31, 32, 33, 34, 35, 38, 39, 40, 41, 42, 43; group V series - 5, 9; Lane 21:  $\lambda$  DNA cleaved with EcoRI + HindIII; Lane 22: uncut pBR322 DNA; Lane 23: linear pBR322 DNA as control. Electrophoresis was carried out on a 1% agarose gel at 4 V/cm overnight.

**Figure 4.** Characterization of plasmid DNA's by double digestion with PvuII and BamHI. DNA's from clones were isolated and digested with two restriction enzymes, PvuII and BamHI. The DNA fragments were analyzed by electrophoresis on an agarose gel (1%) at 4 V/cm overnight. The gel was photographed after staining with ethidium bromide. Lanes 1-9: DNA's from clones I-32, I-44, II-14, II-19, II-43, III-3, III-44, V-28 and V-32; Lane 10:  $\lambda$  DNA cleaved with EcoRI + HindIII; Lane 11: uncut pBR322; Lane 12: pBR322 cleaved with PvuII + BamHI as control.

**Figure 5.** Further characterization of two clones, I-32 and II-19, by cleaving the plasmids with PstI.

The plasmid DNA's from two clones, I-32 and II-19, were purified and cleaved with PstI. The digests were fractionated on an agarose gel (1%). The sizes of the inserts released from the plasmids were estimated to be approximately 400 bp. Lanes 1 and 3 contained uncleaved DNA's from I-32 and II-19 clones. Lanes 2 and 4 contained PstI digest of clones I-32 and II-19. The DNA insert from I-32 (pNK132) was found to be slightly longer than the insert from II-19 clone and hence was chosen for further characterization.

**Figure 6.** Size analysis of cDNA synthesized from Dengue RNA using nicked pNK132 as primer.

pNK132 DNA was digested with PstI, and the cDNA insert was purified by agarose gel electrophoresis. The 400 bp insert (0.1  $\mu$ g) was nick-translated in a reaction mixture containing, in a total volume of 50  $\mu$ l, 50 mM Tris-HCl, pH 7.5, 10 mM MgCl<sub>2</sub>, 1 mM dithiothreitol (DTT), 50  $\mu$ g/ml BSA, 1 mM each of dATP, dGTP and dTTP, 200  $\mu$ Ci of  $\alpha$ -<sup>32</sup>P-dCTP (800 Ci/mmol), 0.3 ng of pancreatic DNase I, 22 units of DNA polymerase I. The reaction mixture was incubated at 100° for 5 min and then chilled quickly in ice. The nicked DNA fragments were used as primers in a reverse-transcriptase-catalyzed reaction containing 0.15 pmoles of Dengue RNA under the same conditions described in the text for cDNA synthesis. The cDNA was purified by Sephadex G-75 column chromatography. Size analysis was carried out by alkaline agarose gel electrophoresis.

**Figure 7.** Procedures for M13 cloning and Sanger's dideoxy DNA sequence analysis.

**Figure 8.** M13mp8 vector used for M13 cloning.

**Figure 9.** Autoradiography of short oligodeoxynucleotides from calf thymus DNA for use in shotgun cloning of Dengue RNA.

Calf thymus DNA (250 mg) was digested with pancreatic DNase I (total volume 7 ml). After incubating at 37°C for 60 min, the reaction was stopped by the addition of 3 ml of 10% SDS and 1.5 ml of pronase (20 mg/ml). The reaction mixture was further incubated at 37°C for 45 min. The solution was extracted once with phenol and 3 times with diethyl ether, and then the deoxynucleotide mixture was precipitated with 2.5 volumes of ethanol in the presence of 0.2 M NaCl. The size of the oligonucleotides was determined by a 20% polyacrylamide gel electrophoresis using markers of known length.

**Figure 10.** Size analysis of cDNA synthesized using random primers on the DEN-2 RNA template.

The reverse transcriptase reaction was carried out using 0.35 pmoles of DEN-2 RNA, 0.12  $\mu$  moles of calf thymus oligonucleotides and 48 units of the enzyme, either in the presence (lane b) or absence (lane a) of actinomycin D. The size of cDNA was determined by alkaline agarose gel electrophoresis.

**Figure 11.** "Flip-back" mechanism for double-stranded cDNA synthesis.

**Figure 12.** Fractionation by Sepharose 4B column chromatography of double-stranded cDNA synthesized using random primers on DEN-2 RNA template.

Double-stranded cDNA was synthesized by self-priming "flip-back" mechanism using Klenow DNA polymerase L. The incubation was carried out in 0.1 M Hepes, pH 6.9, 20 mM MgCl<sub>2</sub>, 2 mM DTT, 60 mM KCl, 1 mM of each of the deoxynucleoside triphosphates at 15°C for 16 h. After addition of 5' phosphorylated Sall linker using T4-DNA ligase, the dsDNA was digested with S1 nuclease. The cDNA was fractionated by size using Sepharose 4B column chromatography. Fractions were analyzed by agarose gel electrophoresis and by autoradiography.

**Figure 13.** Size analysis of M13 recombinant DNA's from various clones. The M13 DNA's were isolated from colorless plaques obtained from transformation of *E. coli* JM103 cells with recombinant M13mp8 vector DNA. a. Lane 1:  $\lambda$  DNA cleaved with EcoRI + HindIII; Lane 2: M13mp8 vector DNA control (with no insert); Lanes 3-21: clones, SMSL series - 001, 005, 010, 015, 020, 025, 030, 035, 040, 045, 050, 055, 060, 065, 070, 075, 080, 085, 090. b. Lanes 1-3 and 5-12: clones, SMSL - 012, 013, 014, 015, 016, 017, 018, 019, 020, 021 and 022; Lane 4: M13mp8 vector DNA control; Lane 13:  $\lambda$  DNA cleaved with EcoRI + HindIII as marker DNA's. Electrophoresis was carried out on an agarose gel (1%) at 4 V/cm overnight. The gel was stained with ethidium bromide and photographed.

**Figure 14.** Dot-hybridization of M13 DNA's with ( <sup>32</sup>P ) labeled cDNA synthesized from DEN-2 RNA as template and random DNA primers, washed and autoradiographed. The experimental details are described in the text. Twenty-eight clones gave positive signals by this technique. Prolonged exposure of the filter to an x-ray film (for one month) gave darker dots for some of these clones. The differences in the intensities of the dots might be due to the underrepresentation of cDNA molecules synthesized from certain regions of DEN-2 RNA due to secondary structures present in the viral RNA.

**Figure 15.** Size analysis of M13 DNA's isolated from purified plaques. The colorless plaques which gave positive signals in dot-hybridization as well as by size analysis shown in Fig. 13 were further purified by plating on agar plates containing minimal media containing glucose as described in the text. The colorless plaques were picked from these plates and grown. The M13 DNA's were isolated and analyzed by agarose gel electrophoresis (1%) as described in Fig. 13. Lane 1:  $\lambda$  DNA cleaved with EcoRI + HindIII as marker DNA's; Lanes 2-6 and 8-13: clones SMSL series - 005, 006, 008, 009, 010, 012, 015, 016, 018, 023 and 024; Lane 7: M13mp8 vector DNA control.

**Figure 16.** Dot-hybridization of M13 clones to nick-translated 400 bp DNA insert from pNK132 plasmid.

The 400 bp DNA insert was isolated by PstI cleavage of pNK132 DNA, followed by agarose gel electrophoresis. The DNA fragment was nick-translated using ( $\alpha$ - $^{32}$ P) dCTP (800 Ci/mmol), pancreatic DNase I and DNA polymerase I. A total of 79 M13 clones were applied to two nitrocellulose filters and hybridized to the probe. The results are shown in Fig. 16. Six positive clones SMSL-023, 024, 028, 102, 103 and 114 were found to hybridize to the 400 bp DEN-2 cDNA insert.

**Figure 17.** Digestion of 5'-labeled 400 bp PstI-fragment from pNK132 plasmid with AluI, HhaI, HpaII and TaqI.

The 400 bp DNA insert was isolated by agarose gel electrophoresis of PstI digest of pNK132 plasmid which was dephosphorylated with calf intestinal alkaline phosphatase before gel fractionation. The DNA fragment was labeled with ( $^{32}$ P) at the 5' termini using polynucleotide kinase and  $\gamma$ -AT $^{32}$ P. The labeled DNA was purified by Sephadex G-50 column chromatography. An aliquot (0.1  $\mu$ g) of DNA was digested with AluI (Lane 1), HhaI (Lane 2), HpaII (Lane 3), or TaqI (Lane 4). The digests were fractionated by PAGE (20%). The gel was dried and autoradiographed. The reason for the abnormal mobility of DNA in HhaI digest is unknown.

**Figure 18.** TaqI digestion of 5'-labeled PstI-fragment (400 bp) of pNK132 plasmid.

The 5'-labeled DNA fragment (3  $\mu$ g) was digested with TaqI and applied to an 8% PAGE. The gel was autoradiographed for 30 min. The bands A and B were cut out, eluted from the gel and used for sequence analysis by Maxam and Gilbert (1977) method.

**Figure 19.** DNA sequence analysis of TaqI-B fragment of 400 bp, PstI DNA fragment of pNK132 plasmid. Autoradiography of the sequencing gel.

**Figure 20.** DNA sequence analysis of TaqI-B fragment.

**Figure 21.** DNA sequence analysis of an M13 DNA clone by Sanger's dideoxy chain termination method. Autoradiography of the sequencing gel obtained from the clone SMSL-008.

**Figure 22.** DNA sequence data.

DNA sequences were obtained from some of the M13 DNA clones using Sanger's dideoxy method. These are yet to be analyzed by computer using the program developed by Staden (1980).

# FIGURE 1

Sept. 1983  
Annual Report

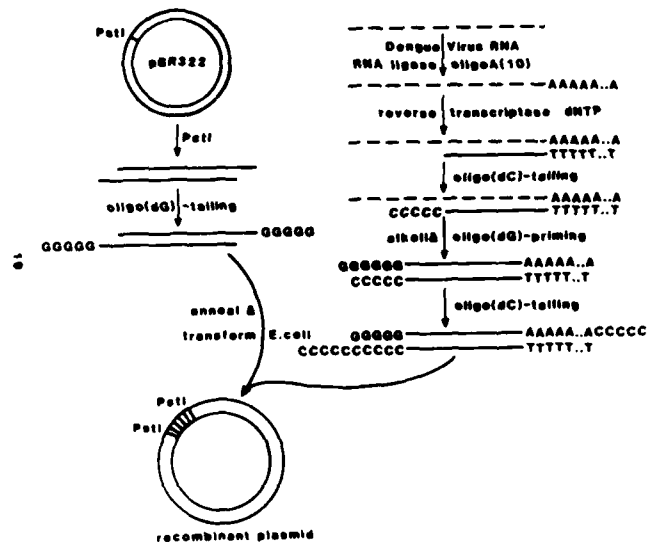


FIGURE 2

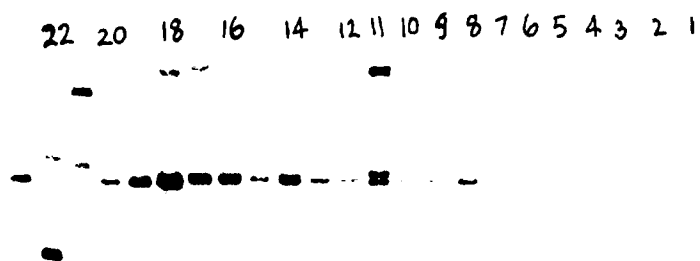
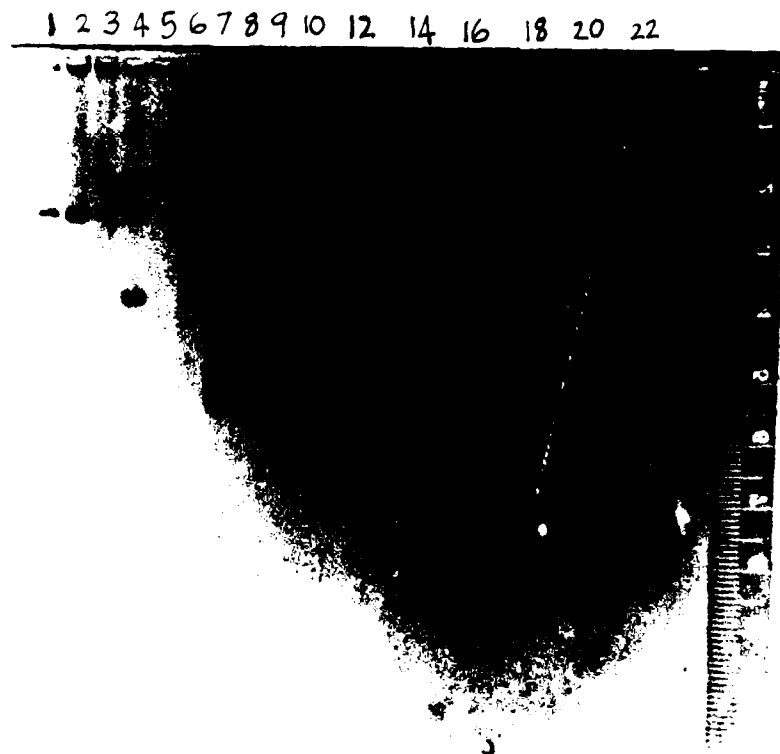


FIGURE 3



### Figure 4

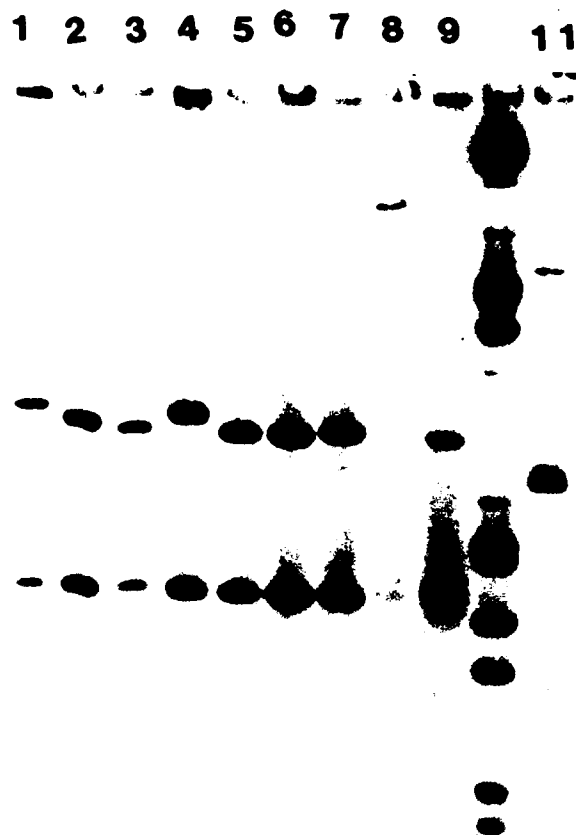


FIGURE 5

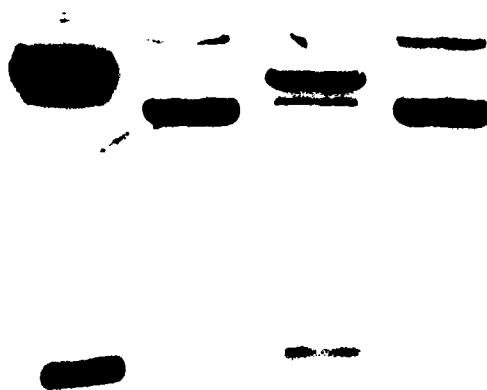
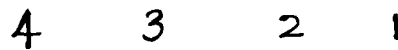


FIGURE 6

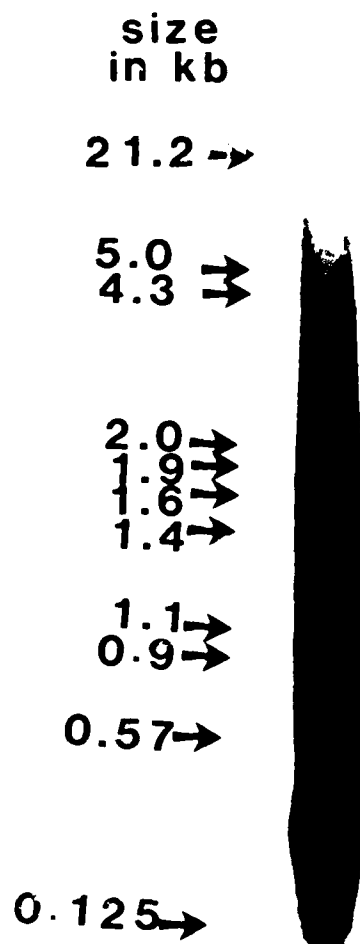
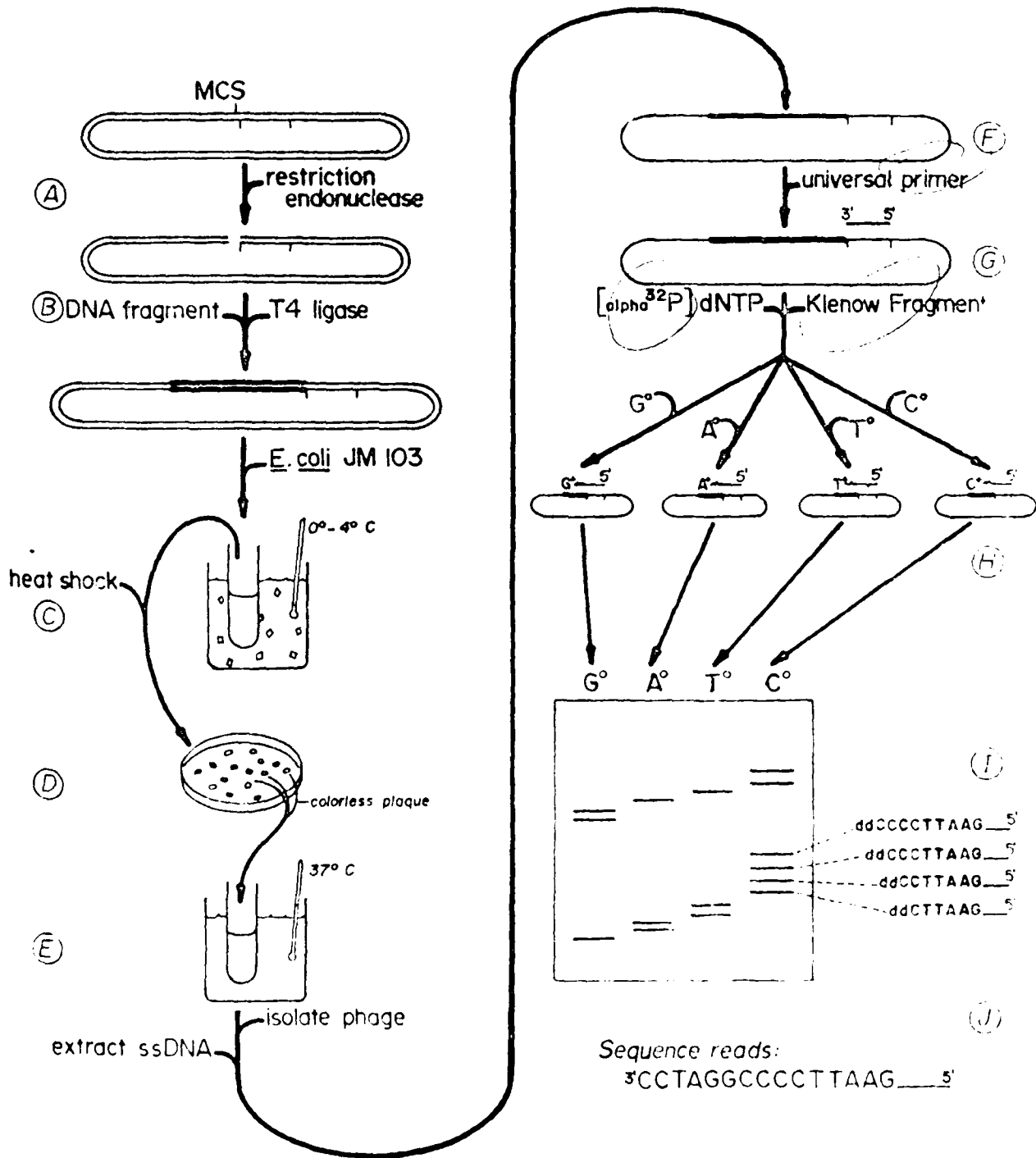
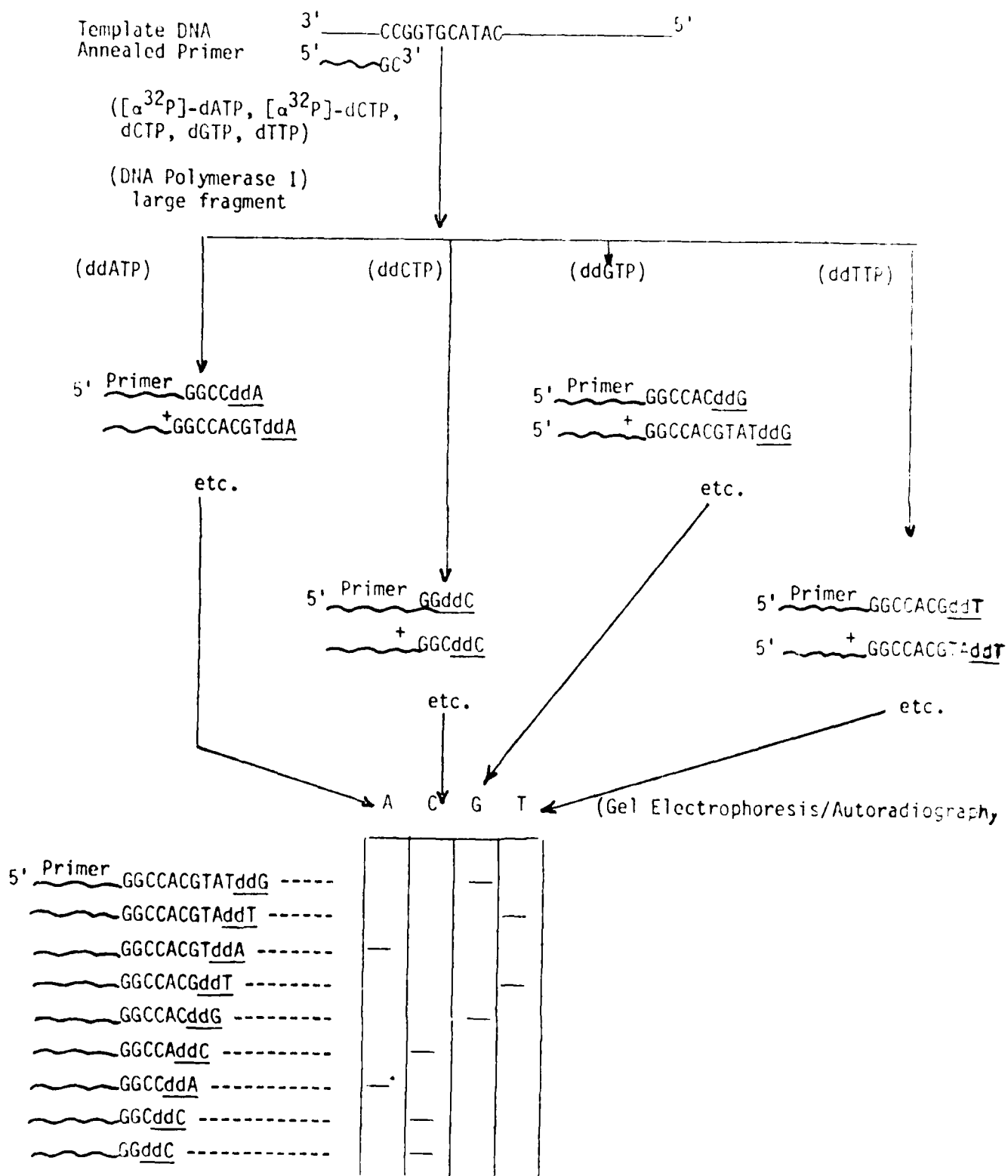


FIGURE 7

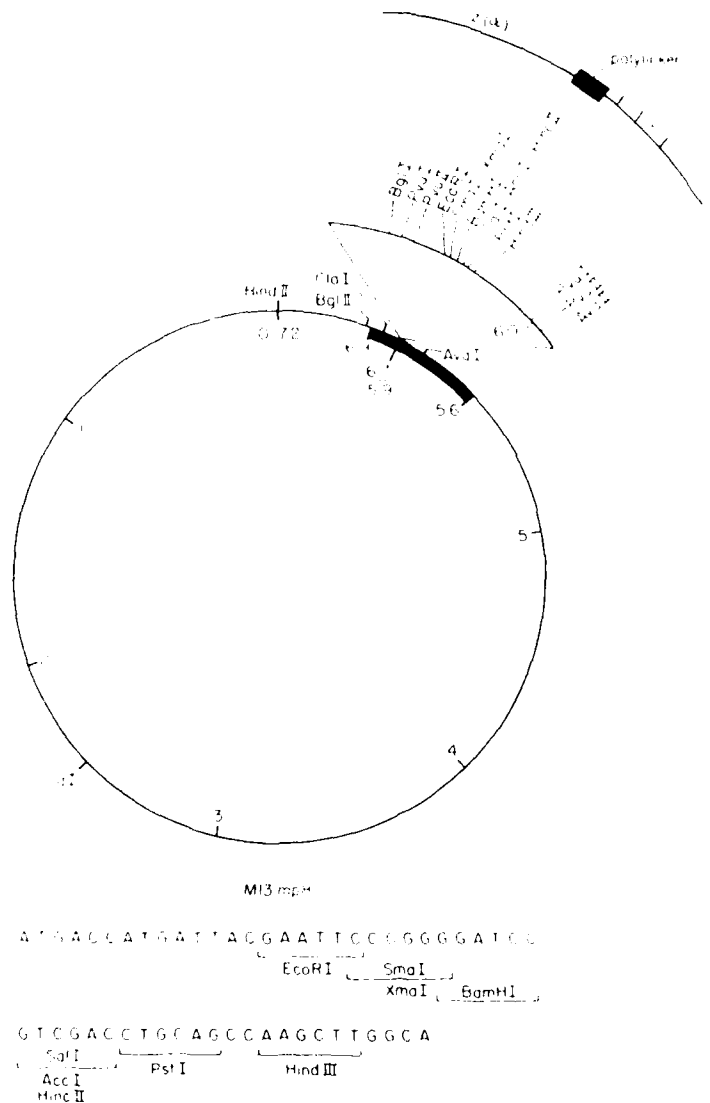


# FIGURE 7 (contd)

## DNA Sequencing with Dideoxynucleoside Triphosphates as Chain Terminators (The Strategy)



# FIGURE 8



1

This M13 cloning vector (mp8), developed by J. Messing, contains a "polylinker" region that makes it possible to clone a variety of restriction-enzyme-generated DNA fragments. Insertion of a foreign DNA fragment into the polylinkers inactivates the  $\beta$ -galactosidase gene (and facilitates a simple screening procedure for phage containing inserts that are not  $\alpha$ -complementable).

FIGURE 9

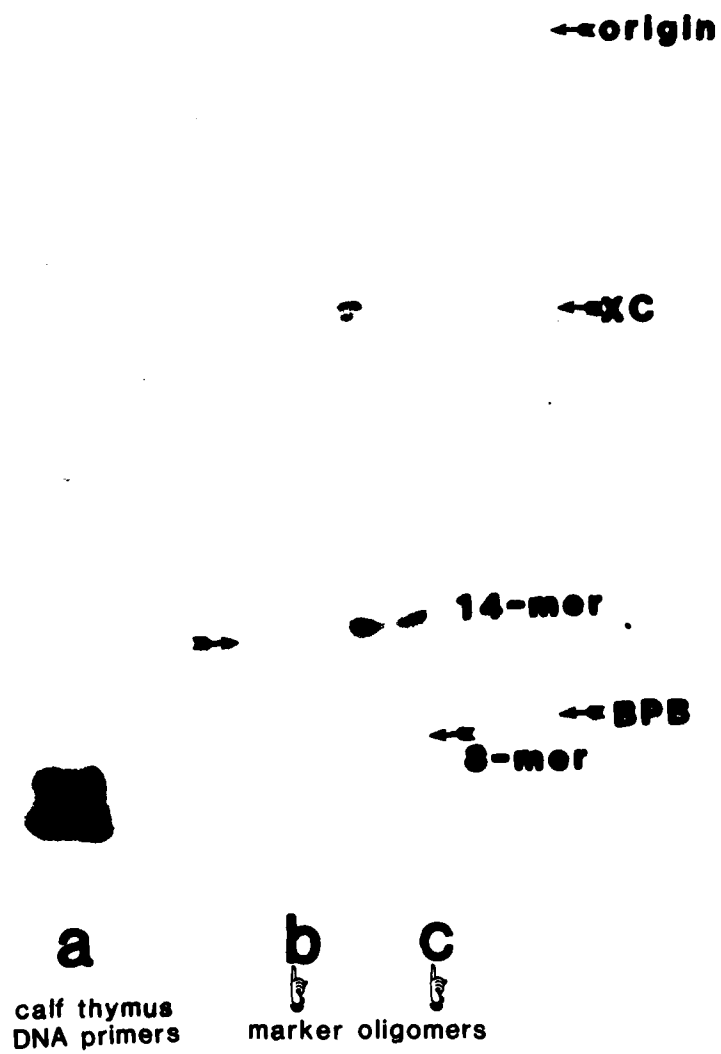
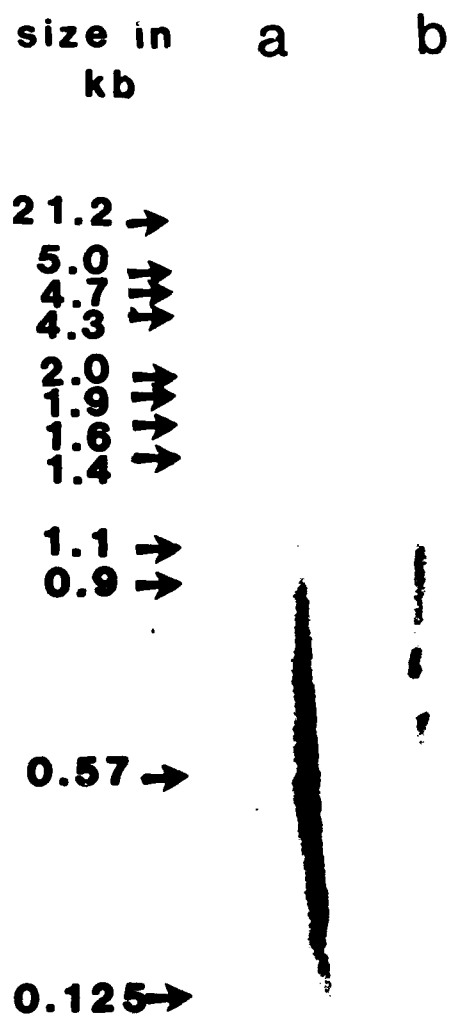
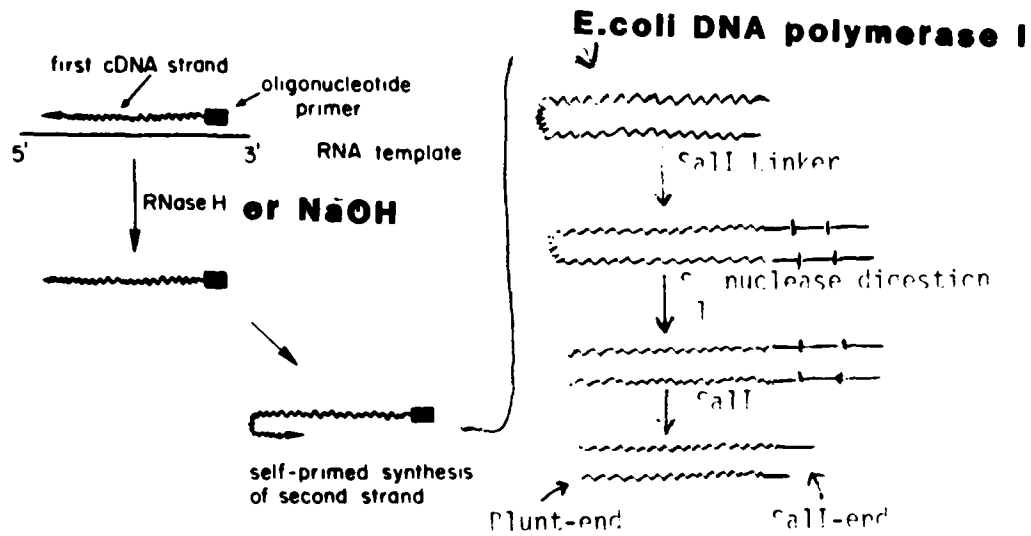


FIGURE 10



**Figure 11**



**FIGURE 12**

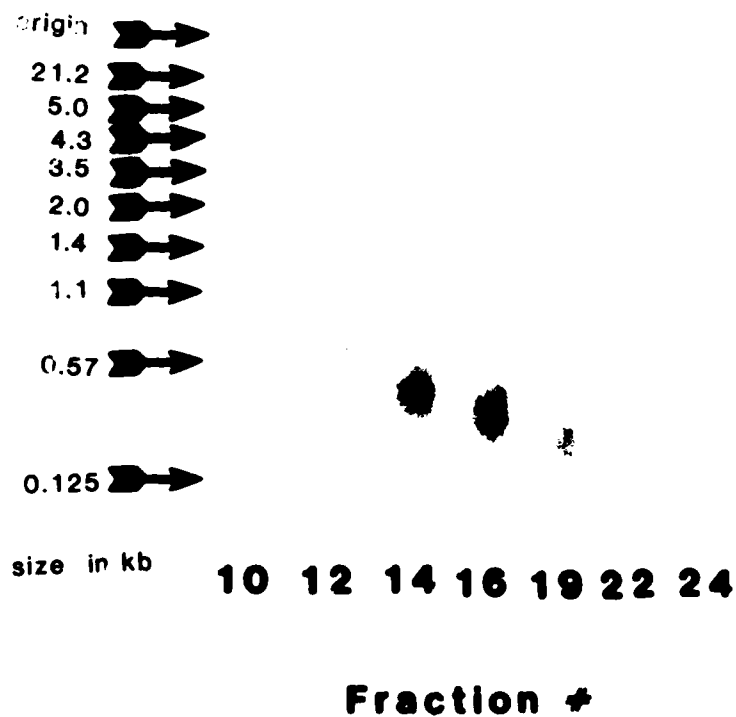
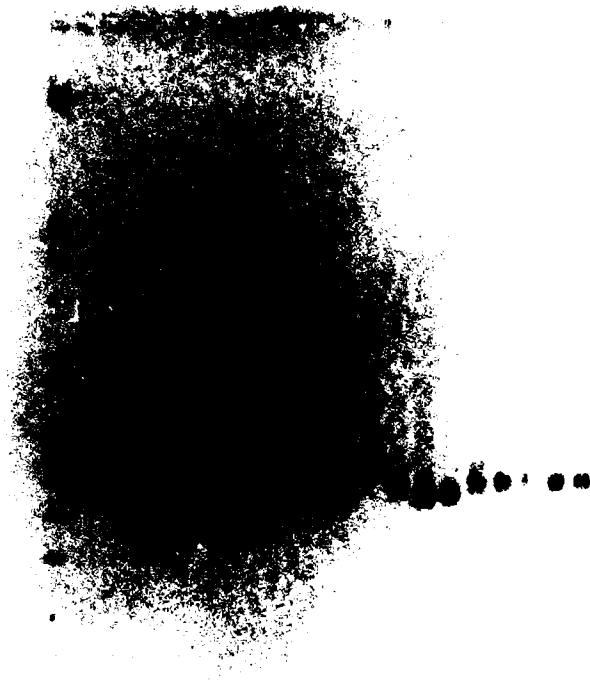


FIGURE 13

a



b

1

13

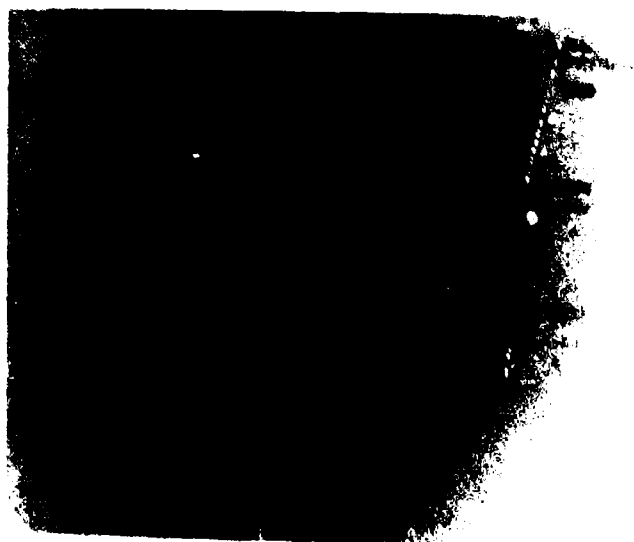


FIGURE 14



FIGURE 15

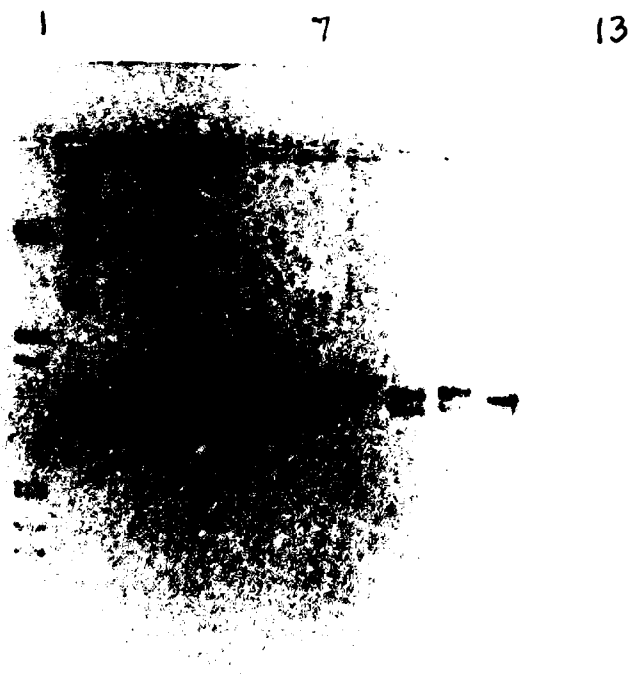


FIGURE 16

SLSM 1023

SLSM 1024

SL SM 10 28

SLSM1103  
SLSM1102  
SLSM 1114

control DNA 1 2 3 4

FIGURE 17



FIGURE 18



FIGURE 19

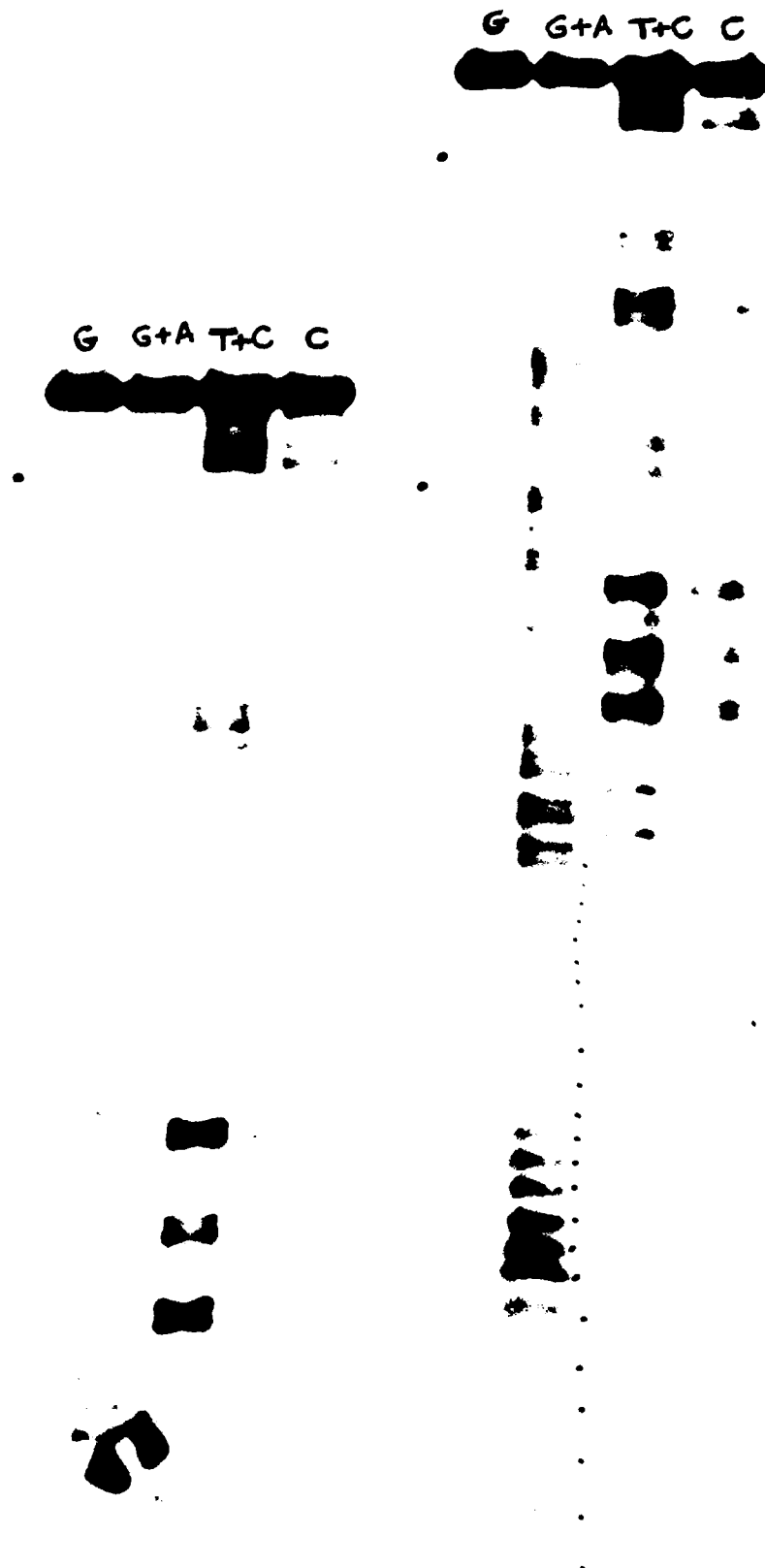


Figure 28

100 base pair sequence of MGNK122

10	20	30	40	50	60
GGGGG GGGGG	GGGGG GGGGG	GGGGG GGGGG	CATAC GATA	TGGGT TATTG	GATAG AAGGT
70	80	90	100	110	120
GCACT CAATG	ACACA TGGAA	GATAG AGAAA	GCCTC TTTC		

FIGURE 21

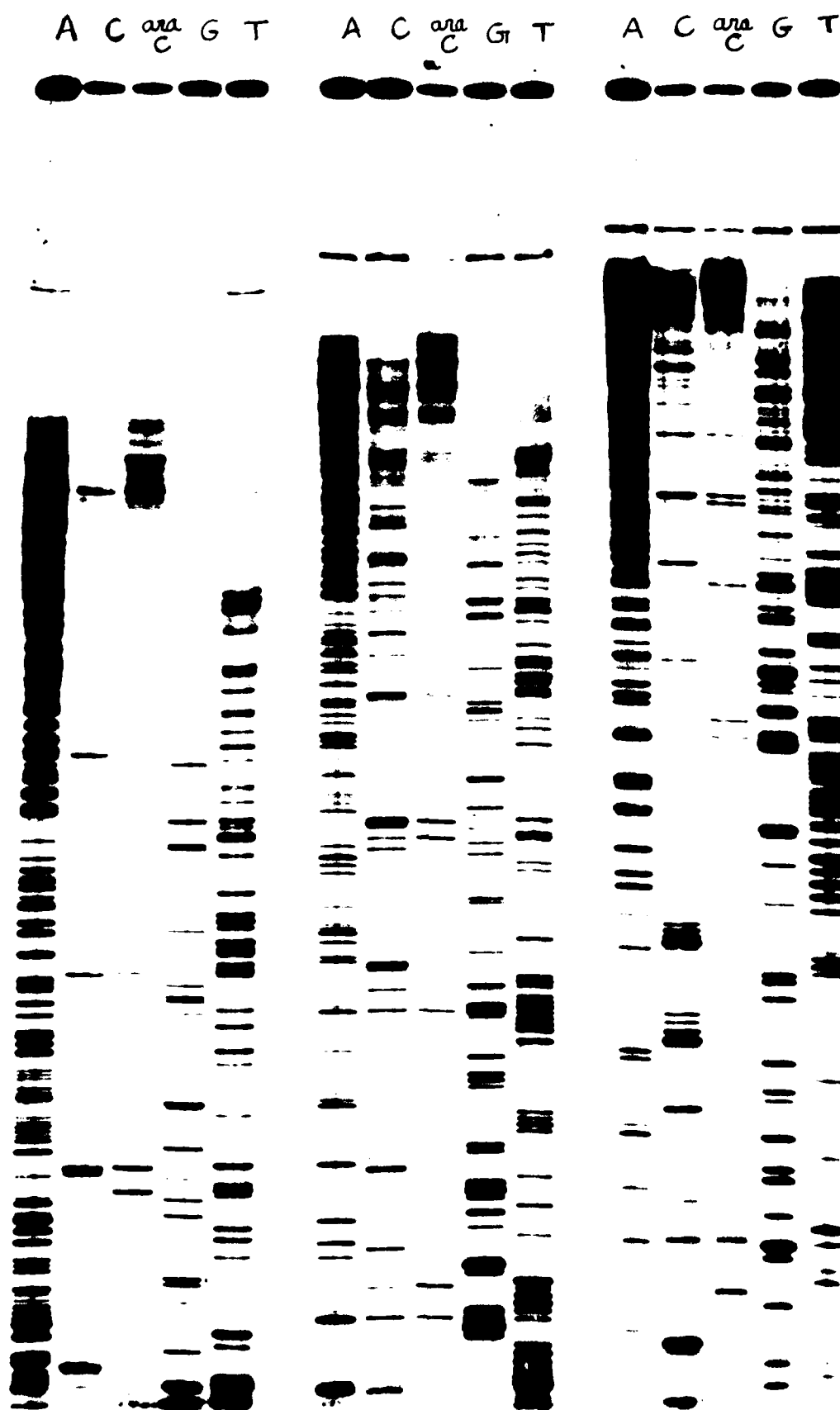


Figure 22

## 132 base pair sequence of SLSM1006

```

10      20      30      40      50
CCCAT TGGAT ATTTT TGGTA ATTTT TAAAC GCATT AATT  TCGAA CTGA  TTTT AA
70      80      90     100     110
ATTAG AGCTT TTTT AAGCA ATTTT ATTCA AAAGG ATTTT TGGTT TAAAT TTTT TAA
130     140     150     160     170
TGATT CCCCC CT

```

## 225 base pair sequence of SLSM1008

```

10      20      30      40      50
CCCCA GAGGT TCCCC CATGT ATATG TACTG GTCAT TTAA  TTTT TGGAG TTTT TAA
70      80      90     100     110
GAGTG GGTCA CGCAT TTATA AAGCA AGACT GTTTA GTTGG TTCCA AGTAT AAAA TAA
130     140     150     160     170
AATGA TATAG CAGCT TGCCT GAGAA AAAAT GGAAA GAAAG TGATA CACCT CAGTA TAA
190     200     210     220     230
ACTTT GATTC TGAAT ATGTC AAGAC TAGAC CAGTA TTGGA CTGCT

```

## 249 base pair sequence of SLSM1009

```

10      20      30      40      50
CGGTC GACCC CCAGA GTTTA CCCCC TOTAT ATGTA CTGCT CATTT TAATT TTTT GATT
70      80      90     100     110
CTGGA CATGA GTGGG TCGAG GATTT TAAAG GGAAG ATGTT TGGTT CTTTC CAAGT ATAAA
130     140     150     160     170
AAGCA GAAAT GATAT AGCAG CTGTC CTGAG AAAA ATGGA AAGAA AGTGA TACAA GTG
190     200     210     220     230
TAGGA AGACT TTTGA TTCTG AGTAT GTACA GACTA GAGCA ATGAT TGGAC TCGT
250     260     270     280     290
CTGCA ATGC

```

## 209 base pair sequence of SLSM1012

```

10      20      30      40      50
TGTAC AACAG CTGAC AAGAA GATTC TCAGT TGGAA TCGTG CAGGG AGGAG GACCA TTAAG
70      80      90     100     110
ACTGT TCATG GCGCT GGTCC CTTCG CTTCG TTTC TAACA ATACC ACCAA CAGCA GTGAT
130     140     150     160     170
ACTGA ACCCA TGGGG AACAA TTAAG ANNNN NNNNG CCATT AATGT TTTGA GAGGT TCAAT
190     200     210     220     230
GAAGG AGATT GGAAC GAGCT GACAC TGAC

```

Figure 22(contd)

## 224 base pair sequence of SLSM1011

```

10      20      30      40      50
TATA TATA TATA TATA TATA TATA TATA TATA TATA TATA TATA TATA TATA TATA
70      80      90     100     110
TATA TATA TATA TATA TATA TATA TATA TATA TATA TATA TATA TATA TATA TATA
130     140     150     160     170
TATA TATA TATA TATA TATA TATA TATA TATA TATA TATA TATA TATA TATA TATA
190     200     210     220     230
TATA TATA TATA TATA TATA TATA TATA TATA TATA TATA TATA TATA TATA TATA

```

## 149 base pair sequence of SLSM1014

```

10      20      30      40      50
TATA TATA TATA TATA TATA TATA TATA TATA TATA TATA TATA TATA TATA TATA
70      80      90     100     110
TATA TATA TATA TATA TATA TATA TATA TATA TATA TATA TATA TATA TATA TATA
130     140     150     160     170
TATA TATA TATA TATA TATA TATA TATA TATA TATA TATA TATA TATA TATA TATA
190     200     210     220     230
TATA TATA TATA TATA TATA TATA TATA TATA TATA TATA TATA TATA TATA TATA

```

## 193 base pair sequence of SLSM1015

```

10      20      30      40      50
TATA TATA TATA TATA TATA TATA TATA TATA TATA TATA TATA TATA TATA TATA
70      80      90     100     110
TATA TATA TATA TATA TATA TATA TATA TATA TATA TATA TATA TATA TATA TATA
130     140     150     160     170
TATA TATA TATA TATA TATA TATA TATA TATA TATA TATA TATA TATA TATA TATA
190     200     210     220     230
TATA TATA TATA TATA TATA TATA TATA TATA TATA TATA TATA TATA TATA TATA

```

## 143 base pair sequence of SLSM1013

```

10      20      30      40      50
TATA TATA TATA TATA TATA TATA TATA TATA TATA TATA TATA TATA TATA TATA
70      80      90     100     110
TATA TATA TATA TATA TATA TATA TATA TATA TATA TATA TATA TATA TATA TATA
130     140     150     160     170
TATA TATA TATA TATA TATA TATA TATA TATA TATA TATA TATA TATA TATA TATA
190     200     210     220     230
TATA TATA TATA TATA TATA TATA TATA TATA TATA TATA TATA TATA TATA TATA

```

### Literature Cited

- Birnboim, H. C. and J. Doly (1979) A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucl. Acids Res.* 7, 1513-1523.
- Blair, C. D. and Schmaljohn, C. S. (1979) Characterization of Japanese Encephalitis Virus RNA. Annual Meeting of American Society of Microbiology, Abstracts, Page 300.
- Cann, A. J., et al. (1983) Poliovirus type 3 molecular cloning of the genome and nucleotide sequence of the region encoding the protease and polymerase proteins. *Nucl. Acids Res.* 11, 1267-1281.
- England, T. E. and Uhlenbeck, O. C. (1978) 3'-terminal labeling of RNA with T4 RNA ligase. *Nature* 275, 560-561.
- Goelet, P., Lomonosoff, G. P., Butler, P. J. G., Akam, M. E., Gait, M. J. and Karn, J. (1982) Nucleotide sequence of tobacco mosaic virus RNA. *Proc. Nat. Acad. Sci. USA* 79, 5818-5822.
- Maniatis, T., Fritsch, E. F. and Sambrook, J. (1982) Molecular cloning: A laboratory manual (Cold Spring Harbor Laboratory Press) Cold Spring Harbor, N.Y.
- Maxam, A. M. and Gilbert, W. (1977) A new method for sequencing DNA. *Proc. Nat. Acad. Sci. USA* 74, 560-564.
- Padmanabhan, R., Jay, R. and Wu, R. (1974) Nucleotide Sequence analysis of DNA XV. Chemical synthesis and sequence analysis of a segment of lysozyme gene for bacteriophage T4. *Proc. Nat. Acad. Sci. USA* 71, 2510-2514.
- Padmanabhan, R. (1977) Chemical synthesis of an octanucleotide complementary to a portion of the cohesive end of P2 DNA and studies on the stability of duplex formation with P2 DNA. *Biochemistry* 16, 1996-2003.
- Petterson, B. (1979) Preparative and analytical purification of DNA from agarose. *Proc. Natl. Acad. Sci. USA* 76, 615-619.
- Sanger, F., Nicklen, S. and Coulson, A. R. (1977) DNA sequencing with chain terminating inhibitors. *Proc. Nat. Acad. Sci. USA* 74, 5463-5467.
- Sanger, F., Coulson, A. R., Barrell, B. G., Smith, A. J. H. and Roe, B. A. (1980) Cloning in single-stranded bacteriophage as an aid to rapid DNA sequencing. *J. Mol. Biol.* 143, 161-178.
- Staden, R. (1980) A new computer method for the storage and manipulation of DNA gel reading data. *Nucl. Acids Res.* 8, 3673-3694.
- Taylor, J. M., Illmensee, R. and Summers, J. (1976) Efficient transcription of RNA into DNA by avian sarcoma virus DNA polymerase. *Biochim. Biophys. Acta* 442, 324-330.

Personnel Supported from September 1, 1982-August 31, 1982

<u>Personnel</u>	<u>Title</u>	<u>Period Supported by Contract</u>
Dr. Narender Kalyan	Research Associate	September 1, 1982-present
Dr. Masatoshi Nohara	Research Associate	September 1, 1982-June 30, 1983
Maj. John R. Lowe*	Graduate Student	September 1, 1982-present
Parvesh Kumar	Medical Student	January 1983-August 17, 1983
Ronnie Pelton	Summer Undergraduate Student	June 1, 1983-July 31, 1983
Steve Copenhaver	Summer Undergraduate Student	June 1, 1983-July 31, 1983
Baljit Kaur	Dishwasher	March 1983-present
Dr. Radha Krishnan Padmanabhan	Principal Investigator	September 1, 1982-present

\*Research materials and Institutional overhead support only. Academic and salary support is from USAMEDD Long Term Civilian Training Program.

DATE  
FILMED  
-8